HAVE PERFORMANCE REQUIREMENTS HISTORICALLY BEEN MET IN SYSTEMS DEVELOPED FOR THE U. S. MILITARY?

Dr. Edmund H. Conrow, Certified Management Consultant
12 June 1997
Tel: (310) 374-7975
Email: ehc@earthlink.net
Web Site: http://home.earthlink.net/~ehc/

1.0 INTRODUCTION

During the SCEA Acquisition Reform Model Sharing (ARMS) Workshop on 10 June 1997 a question was raised about the characteristics of the database I analyzed to estimate the extent that performance requirements have been historically been met in developing major DoD systems. Specifically, for military development programs is the required level of performance met, **on average**, or is reduced performance common as is cost growth and schedule slippage?

A hypothesis I advanced at the ARMS Workshop is that for most major military development programs performance requirements are met **on average** while cost and/or schedule are often adjusted (typically upward) in order to meet performance. I will now present evidence from other studies, as well as my own study, that strongly supports this hypothesis. (For those just interested in the "bottom line", see Section 3.0.)

I will now briefly digress to discuss how changes in cost, performance and schedule are typically reported in the literature, because this concept is needed for the reader to understand the results presented in Sections 2 and 3. The most common reporting method is known as the change ratio. It is computed by taking the ratio of a variate (e.g., performance) achieved versus that originally specified (required). Values > 1 indicate the measured level of performance is worse than required, values = 1 indicate a performance level equal to that required, while values < 1 indicate a performance level better than required. In terms of cost and schedule change, values > 1 are often termed "growth" or "slippage" and are not desirable, while values < 1 are desirable. For performance change, two conventions were established by Perry, et. al. in 1971 [1] that are still followed today. First, the average of performance values is computed for a given program, since

all programs will have more than a single measure of required performance. Second, performance attributes that have characteristics which are desirable with increasing levels are inverted so that the resulting ratio will be < 1 for (beneficial) improvements over the required level. (For a fuller discussion of change ratios and their computation, see [2].)

For example, if the required missile speed is Mach 1.0 and the measured missile speed is Mach 1.25, the resulting performance change ratio in this case is 1.0/1.25 = 0.80. Similarly, if the required missile weight is 1000 pounds and the measured missile weight is 950 pounds, the performance change is 950/1000 = 0.95. The average performance change for the hypothetical missile program, based upon the two example attributes here, is thus (0.80 + 0.95)/2 = 0.875 (rounded to 0.88).

2.0 STUDY RESULTS

2.1 Marshall and Meckling Study

Perhaps the earliest discussion of why performance degradation is uncommon, while cost and schedule slippage are common in the development of military systems is given by Marshall and Meckling of The Rand Corporation in 1959:

> "Typically, in weapons development great emphasis is placed on performance. Most new weapons are developed around specific detailed performance requirements laid down by the military--requirements that are taken very seriously. The penalties incurred by the contractors for not meeting performance requirements are more severe than for failure to meet availability schedules or failure to live within original cost estimates. As a result, whenever circumstances dictate a retreat from early plans, it is usually the costs and/or availability that gives ground. If for example, technological difficulties with electronics equipment prohibits the attainment of the required performance on the original schedule at original costs, usually either the availability date is postponed, or costs are increased or both. Degradations in performance are seldom tolerated." [3]

Marshall and Meckling included an analysis of cost and schedule change during the military development process (both cost and schedule typically had considerable growth on most programs), but did not statistically analyze how performance varied. So while Marshall and Meckling predicted that performance change would likely be near 1.0 ("degradations in performance are seldom tolerated") while cost and schedule growth would typically occur, they did not prove that minimal performance degradations were the rule.

2.2 Perry et. al. Study

A study performed by Perry, et. al. of The Rand Corporation in 1971 was the first published study that assessed how cost, performance and schedule varied during the course of developing and

producing major military development programs [1]. The Perry data set was originally derived from surveys and follow-up visits to various DoD programs: it represents programs with Milestone II (or equivalent) dates in the 1950s (2 programs) and 1960s (18 programs). While some of the programs examined were in engineering development, many were mature and in the production phase. In addition, some cost, performance and schedule change values existed in this database that were estimates or collected before the program's Initial Operational Capability (IOC) date.

Perry found that **on average** the level of performance achieved versus that specified at the start of engineering development (equivalent to Engineering and Manufacturing Development (EMD) today) was 1.05--a five percent slip in performance. (When a few errors were removed from this database, the resulting performance change was 1.03, as reported in Section 2.4.) For cost and schedule change, the average computed ratios were 1.44 and 1.15, respectively [4]. The Perry study also concluded the following:

> "First, **on average**, performance objectives were met while schedules were allowed to slip somewhat and costs even more. Also, it is clear from the entire distribution of outcomes that performance was frequently attained or exceeded; to a lesser extent the same is true of schedules; but costs have invariably been higher than predicted." [5]

Both the Perry results and the above statement strongly support the hypothesis that I advanced at the ARMS Workshop--that **on average**, performance requirements are typically met while cost and/or schedule are often adjusted (typically upward) during the course of acquiring major military programs.

2.3 Dews et. al. Study

A study performed by Dews, et. al. of The Rand Corporation in 1978 [6] extended the work of Perry et. al. in examining how cost, performance and schedule varied during the course of developing and producing major military development programs. The Dews data set was derived from Selected Acquisition Reports (SARs), and represents programs with Milestone II (or equivalent) dates in the 1970s. While some of the programs examined were in engineering development, others were mature programs in the production phase. In addition, some cost, performance and schedule change values existed in this database that were estimates or collected before the program's IOC date.

Dews found that **on average** the level of performance achieved versus that specified at the start of full scale development (equivalent to EMD today) was 1.00--performance exactly meeting requirements. For cost and schedule change the average computed ratios were 1.34 and 1.13, respectively [7].

The Dews results strongly support the hypothesis that I advanced at the ARMS Workshop--that **on average**, performance requirements are typically met while cost and/or schedule are often adjusted (typically upward) during the course of acquiring major military programs.

2.4 Conrow Study

Collection and analysis of data for the Conrow study began in 1981 and was substantially aided by Robert Perry (Perry et. al. primary author), Edmund Dews (Dews et. al. primary author), Giles Smith (Perry et. al. and Dews et. al. co-author) and Allen Barbour (Dews et. al. co-author) both while I was at The Rand Corporation (1981-1983) as well as after I left through 1994--a 13 year period of time. Giles Smith provided considerable assistance from 1981 to 1994 in evaluating this data, and Smith, Edmund Dews and Andrew Marshall reviewed my results [8]. Of course, any errors in the computation of results should be solely attributed to Conrow.

The methodology used by Conrow was virtually identical to that used in the Perry and Dews studies--evolving a bit to reduce potential errors and to make it more consistent across individual data sets. Data from the Perry and Dews databases were re-constructed and screened to evaluate each program. In a few instances computational errors were corrected which led to slight variations in results.

The Perry schedule change data was available, re-evaluated and used in the Conrow study database. (The re-computed Perry data set schedule change value was 1.14 versus 1.15 reported in the Perry study [4].) The Perry performance change data was available, re-evaluated and used in the Conrow database. (The re-computed Perry data set performance change value was 1.03 versus the 1.05 reported in the Perry study [4].) Unfortunately, the Perry cost change data could not be reconstructed, so only that Perry data reported in the Dews report could be used [9].

Cost, performance and schedule change results from several programs in the Dews database were estimates or actual values obtained prior to completion of the development phase. Conrow re-collected data for a number of these programs from updated SARs, since several of the programs in the Dews database reached IOC after the Dew report was issued in 1978. (Although the IOC date is typically somewhat after the close of program development (Milestone III starts the rate production phase), this milestone was used for cost, performance and schedule change data computation as an indicator of the close of system development since it was not possible to identify the Milestone III date or production start in some cases.)

Two types of corrections were made to the Dews data by Conrow. First, a single schedule milestone (IOC date) was used in computing schedule change, versus multiple schedule milestones used by Dews. This was done in order to make the schedule change data more compatible with cost change data where a single cost change ratio is reported. Second, Dews bounded performance attribute results for each program by 0.5 to 2.5, meaning that any individual performance parameter with an [estimated or actual]/specified performance level ratio > 2.5 or < 0.5 was omitted. This was done to preclude outliers from swamping the average performance change value for a given program. Conrow adjusted the lower bound of this ratio to 0.4 to make

the resulting lower and upper bounds symmetrical (in effect, a factor of 1/2.5 to 2.5/1). These same two adjustments were also made to the Perry data as well as the data collected by Conrow from SARs.

The Conrow database included only results obtained from the first SAR either on or after the program had reached IOC. Consequently, the Conrow data included only change data derived from actual data following completion of the development phase, versus data prior to completion of the development phase in some Perry and Dews database programs. The Conrow data set was derived from SARs, and represents programs with Milestone II (or equivalent) dates in the 1960s (7 programs), 1970s (17 programs), and 1980s (6 programs).

The re-computed Perry and Dews data were then added to the Conrow data to form one overall, combined database. (This was permissible since the statistical characteristics of the Perry, Dews and Conrow databases were similar.) The resulting overall database included cost, performance and schedule change results from the: (1) Perry database--11, 17 and 19 programs, respectively; (2) Dews database--7, 5 and 2 programs, respectively; and (3) Conrow collected data--30, 30 and 30 programs respectively. This yields 48, 52 and 51 cost change, performance change and schedule change major program data points, respectively in the overall database.

Results of the Conrow study were initially reported at the 1995 Defense Systems Management College Acquisition Research Symposium [8]. An enhanced set of results were then reported in the Defense Acquisition University publication **Acquisition Review Quarterly** [2].

Using the overall database, Conrow found that **on average** the level of performance achieved versus that specified at the start of EMD (or equivalent) was 1.00--performance exactly meeting requirements. For cost and schedule the average computed ratios were 1.26 and 1.24, respectively [8] [2]. Furthermore, the median of the overall database was 1.00 for performance change, 1.16 for cost change and 1.13 for schedule change, respectively. This coupled with the skewness coefficient computed for each variable indicated that the performance change distribution was highly symmetrical--as if performance requirements were targeted to be met, while the cost change and schedule change distributions were right-skewed, or allowed to adjust.

Results from the overall database, which spanned a 28 year period (Milestone II or equivalent dates of 1958 to 1986), strongly support the hypothesis that I advanced at the ARMS Workshop--that **on average**, performance requirements are typically met while cost and/or schedule are often adjusted (typically upward) during the course of acquiring major military programs. Furthermore, these results did not change significantly when regressed to the calendar date of the program's Milestone II (or equivalent)--thus no substantial change in the results occurred during a 28 year period of time.

3.0 DISCUSSION

Results from three separate comprehensive studies (Perry [1], Dews [6] and Conrow [8]) plus indications from a fourth study (Marshall and Meckling [3]) strongly support the hypothesis that I

advanced at the ARMS Workshop--that **on average**, performance requirements are typically met while cost and/or schedule are often adjusted (typically upward) during the course of acquiring major military programs.

The data analyzed was from major military programs--mostly major development programs. The data included Air Force, Army and Navy programs, and 79 percent of the programs (46 of 58) were aircraft, helicopter and missile programs. The remaining 21 percent of the programs included ship systems (e.g., anti-submarine warfare systems, but no complete ships), ground systems (e.g., armored vehicles) and other systems. The data set did not include any space programs, but almost all of the 58 total programs had a high level of required performance and associated technology. No follow-on or upgrade programs were included to prevent biasing the results through a potentially incremental development activity. In addition, no programs were included that had gone through EMD, as well as reaching IOC, under Clinton Administration DoD acquisition reform. Several more years will be needed to develop a sample with sufficient size to determine if a statistically significant variation in cost, performance and schedule change has occurred versus the historical data analyzed in the cited studies.

Finally, the results of the studies presented in Section 2 should not be taken to universally apply exactly to all military development programs, rather they apply **on average** to military development programs. Exceptions existed where the average level of performance change for a program was clearly not very close to 1.0, thus indicating a degradation of the average level of performance met versus required for that program. Furthermore, it was common that one or more performance requirements were not met out of the typically 5 to 20 requirements for each program analyzed. However, when performance change was computed across all programs, **on average** the required level of performance was met (average value of 1.00). In addition, 42 of the 52 programs (or 80%) in the overall data set had a performance change < 1.10--indicating **at worst** a 10% performance degradation across these programs for 80 percent of the total sample size. The same cannot be said for cost change and schedule change where the 80th percentile value corresponded to 1.48 (or cost growth of < 48 percent) and 1.50 (or schedule slippage < 50 percent), respectively. These results, coupled with others presented in Section 2, strongly support the hypothesis that **on average** performance requirements are typically met while cost and/or schedule are often adjusted (typically upward) during the course of acquiring major military programs.

4.0 REFERENCES

[1] Robert Perry, Giles Smith, Alvin Harman, and Susan Henrichsen, “System Acquisition Strategies”, The Rand Corporation, R-733-PR/ARPA, June 1971.

[2] Edmund H. Conrow, “Some Long-Term Issues and Impediments Affecting Military Systems Acquisition Reform,” **Acquisition Review Quarterly**, Defense Acquisition University, Summer 1995 (refereed journal).

[3] Andrew W. Marshall and William H. Meckling, “Predictability of the Costs, Time, and Success of Development”, The Rand Corporation, P-1821, October, 1959, pp. 20-21.

[4] Robert Perry, Giles Smith, Alvin Harman, and Susan Henrichsen, op. cit., R-733-PR/ARPA, June 1971, pg. 8.

[5] Robert Perry, Giles Smith, Alvin Harman, and Susan Henrichsen, op. cit., R-733-PR/ARPA, June 1971, pg. 9.

[6] Edmund Dews, Giles Smith, Allen Barbour, Elwyn Harris, and Michael Hesse, “Acquisition Policy Effectiveness: Department of Defense Experience in the 1970s”, The Rand Corporation, R-2516-DR&E, October 1979.

[7] Edmund Dews, Giles Smith, Allen Barbour, Elwyn Harris, and Michael Hesse, op. cit., The Rand Corporation, R-2516-DR&E, October 1979, pg. 57.

[8] Edmund H. Conrow, “Some Long-Term Issues and Impediments Affecting Systems Acquisition Reform,” presented at the 1995 (bi-annual) Acquisition Research Symposium, and published in the “1995 Acquisition Research Symposium Proceedings,” Defense Systems Management College (DSMC), June 1995 (sole author). This paper was co-winner of the David D. Acker award for "Skill in Communications" (best paper of the symposium).

[9] Edmund Dews, Giles Smith, Allen Barbour, Elwyn Harris, and Michael Hesse, op. cit., The Rand Corporation, R-2516-DR&E, October 1979, pg. 51.